

# 指導と評価の一体化へのアプローチ

## — パフォーマンス・モデルとルーブリックの活用 —

An Approach to the Integration of Teaching and Assessment  
— Through the Application of Performance Models and a Rubric —

金子 義隆  
KANEKO Yoshitaka

### Abstract

There are some areas that need to be modified in English Education in Japan. One of those areas is assessment. The Ministry of Education, Culture, Sports, Science, and Technology (MEXT) revised the Courses of Study for elementary and secondary schools for the first time in almost ten years. Along with its revision, MEXT hopes to improve assessment in the English classroom by integrating teaching and assessment. This paper presents a way of integrating teaching and assessment, taking into consideration English lessons of the second year of junior high school. First, the paper discusses the background of the revisions of the Courses of Study and the importance of CAN-DO based goal setting. In addition, it discusses the relationship between teaching and assessment and four significant elements of appropriate assessment. Finally, the paper exemplifies how to set a unit goal and decide on an appropriate performance test based on the four elements above by drawing up performance models and a rubric consisting of criteria which have been apparent through making performance models.

キーワード：指導と評価の一体化 逆向きデザイン パフォーマンス評価 ルーブリック

## 1. はじめに

情報基盤社会、グローバル社会、少子高齢化社会などと言われるようになり、これからの社会は過去に人類が経験しなかった変化や課題に向き合わなくてはならない「予測困難な時代」を迎えようとしている。第四次産業革命の主翼を担う人工知能（AI）の発達により、現存する職業の多くがなくなると言われている。ニューヨーク市立大学教授のキャシー・デビッドソンは、「2011 年度にアメリカの小学校に入学した子供たちの 65% は、大学卒業後には存在していない職業に就くだろう」と予測をしている（文部科学省、2014 年、p.1 から引用）。今

後、AI や IoT の発展により Society 5.0 と呼ばれる社会の到来が予測される中、変化する 10 年後、20 年後の未来を見据えながら「学校」は子どもたちを未来社会の新たな担い手とするべく教育をおこなう大きな責任を負っている。このような状況下において、外国語教育、特に英語教育の責務は大きい。グローバル社会の到来は、さまざまな人種間のコミュニケーションの必要性を意味している。英語をはじめとした語学力は予測困難な時代を切り開いていくためのパスポートとなりうる。

文部科学省（2017 年、2018 年）は、これまでの教育の反省も踏まえつつ新しい時代に対応すべく、全校種の学習指導要領を「学びの地図」と新たに呼び、学校、家庭、地域が共有できるものとなるよう

に期待しつつ、以下6つの観点に整理して改訂をおこなった。

- ① 「何ができるようになるか」(育成を目指す資質・能力)
- ② 「何を学ぶか」(教科等を学ぶ意義と、教科等間・学校段階間のつながりを踏まえた教育課程の編成)
- ③ 「どのように学ぶか」(各教科等の指導計画の作成と実施、学習・指導の改善・充実)
- ④ 「子供一人ひとりの発達をどのように支援するか」(子供の発達を踏まえた指導)
- ⑤ 「何が身についたか」(学習評価の充実)
- ⑥ 「実施するために何が必要か」(学習指導要領等の理念を実現するために必要な方策)

(文部科学省, 2018 年, p.3)

上記の6つの観点の①「何ができるようになるか」とは「目標」のことであり、⑤「何が身についたか」は「評価」に関してである。この両者は、目指す目標とその目標達成状況の評価という観点から表裏一体の事案であり、ひいては②「何を学ぶか」に大きな影響を与えるものである。つまり、指導の始めとして目標を設定し、その目標を達成するための指導を行い、その指導の結果として目標の達成状況の把握を行うことは、まさに「指導と評価の一体化」を意味するものである。

「指導と評価の一体化」を徹底すべく、「外国語能力の向上に関する検討会」(2011 年)は「国際共通語としての英語力向上のための5つの提言と具体的施策」を示した。その「提言1」の中で、生徒に求められる英語力について、その達成状況を把握・検証するよう以下のように提示した。

多くの学校では、学習指導要領に基づく授業が行われている一方、一部の学校では、文法・訳読中心の授業、高校入試や大学入試対策に特化

した授業などが行われているとの指摘がある。中・高等学校では、各学校が、学習指導要領に基づき、生徒に求められる英語力を達成するための学習到達目標を「CAN-DO リスト」の形で具体的に設定することにより、学習指導要領の内容を踏まえた指導方法や評価方法の工夫・改善が容易になる。また、各学校が、学習指導要領の目標を地域の実態や生徒の能力に応じて具体的な目標に設定し直すことにより、すべての子どもたちの英語力の水準向上に資するだけでなく、グローバル社会に通用するより高度な英語力の習得を目指すことも可能となる。

(外国語能力の向上に関する検討会, 2011 年, p.7)

このように、学習指導要領に基づき、各中・高等学校が生徒に求められる英語力を達成するための目標(学習到達目標)を「言語を用いて何ができるか」という観点から、「CAN-DO リスト」の形で具体的に設定することについて提言がなされた。

ここで「学習到達目標」とは何かを確認したい。「各中・高等学校の外国語教育における『CAN-DO リスト』の形で学習到達目標設定のための手引き」(2013 年)によれば、

学習到達目標とは、各学校において、全ての生徒に求められる外国語能力を達成するためのものである。その上で、目標を超えた伸長がみられる生徒については、更にその先を見通すことができ、学習意欲を維持できるような工夫が必要である。例えば、可能であれば、通常の学習活動をより発展、拡充させた内容の活動に取り組むことなどが考えられる。他方、習熟により時間がかかる生徒については、学習意欲を維持し、確実に学習を続けることができるような工夫が必要である。例えば、目標に到達しつつある過程を教員が適切に評価することによって、生徒の更なる学習を支援する工夫が必要であ

る」(p.6)。

## 2. CAN-DO リスト形式の学習到達目標

学習到達目標の意味は上記の通りだが、各学校が「CAN-DO リスト」形式の学習到達目標を設定する目的は何だろうか。前出の「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」によれば、目的は3つある。1つは、生徒が目指す英語力を明確にすることで目標がはっきりすることである。「CAN-DO リスト」の形（指標形式）で目標を明示することで指導の焦点も絞りやすくなり、評価も「～できるかどうか」を測ることとなり、「指導と評価の一体化」につながる。2つ目は、「英語を使って何ができるか」の指標形式の目標の明示により、言語の「形式面」ではなく、「技能・運用面」に重きが置かれることになる。言語は4技能が別々に発達するものではなく、指導すべきものでもない。1つの技能に焦点を当てる指導は、ひいては技能統合的に活用されることとなり、4技能の総合的指導に資することとなる。最後の3つ目は、児童生徒が目標を知ることができることであり、自分の学習に対して主体的に関与していく動機づけとなる。

上述の「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」では、「中・高等学校」のみが記されているが、昨今文部科学省（2014年）は、『今後の英語教育の改善・充実方策について 報告～グローバル化に対応した英語教育改革の五つの提言～』で、小・中・高等学校を通じた一貫した指標形式の目標の設定の必要性をはっきりと示した。つまり、初等中等教育における英語教育において、一貫して「指導と評価の一体化」を求めている。その中で、各小・中・高等学校は各学校の実態に則した「CAN-DO リスト」形式での学習到達目標を設定することが望まれている。そして、目標達成を適切に

測定するための評価の充実も同様に望まれている。

## 3. 指導における評価の目的

評価の充実は非常に大きな課題である。まず、評価を考える前に確認しておきたいこととして、そもそも評価をどう活用するかである。本稿で取り上げる評価とは、学習評価のことである。学習評価には2つ目的がある。1つは、児童生徒の学習習熟度を知り、教材のレベルや活動の強度や頻度、授業の進度、指導法など授業全般を振り返り、教師が指導の改善に生かすためである。生徒の質問に対する応答や授業中の取組みの様子、小テストや英作文など日常の学習活動に対して、意識的、潜在意識的に教師は評価をしている（形成的評価）。これは児童生徒の行動、取組みを評価しているのと同時に、教師は自身の授業を振り返る機会となり、その結果に応じて授業を改善していくこととなる。例えば、ディクテーションを行う際、当初の予定は1度のみ聞かせてから書き取らせるはずであったが、生徒の活動の出来を観察して、想像以上にできていないと分かった場合、もう1度聞かせるとか、2度目は少しスピードを遅くして聞かせるなどして生徒の実態に合わせた指導に変えることが可能となる。教師は絶えず生徒の here and now の情報を活用している。それによって、より効果的な指導・学習が起こることになる。そして、評価の2つ目の目的は、児童生徒の学習の到達度がある一定期間の終了後に評価するためである（総括的評価）。これによって、教師は各児童生徒がある一定期間の学習の目標に対してどの程度達成できたのかを把握することができる。これによって、次の期間や新しいクラスの指導計画の改善につなげることができる。もちろん、児童生徒も自分の学習の成果を振り返る機会となる。このように、学習評価の2つの目的は指導と密接にかかわっていることが明らかである。指導と評価の一体化は、授業の「PDCA サイクル」の構築とも言う

ことができる。

#### 4. テストの備えるべき4つの条件

評価は、前出の「手引き」において、「目標に沿った学習活動を適切に評価できる方法」を選択して実施することが必要とされている（p.12）。つまり、評価において重要であり、その効果を最大化するためには、学習活動を「適切に」評価することであり、そのために適切な評価方法を選ぶことである。評価方法は、測定方法と言い換えることができる。測定方法と言うと、「筆記テスト」が最初に頭に浮かんでくるかもしれないが、それだけではない。他に、面接やプレゼンテーション、エッセイ、ポートフォリオなど多種多様である。

「テスト」と「評価」は同義語として扱われることがあるが、同じではない。「テスト」とは、児童生徒に対して刺激（問題等）を与えて、その反応（解答）を引き出す「方法」のことである。英語の授業で育成しようと目指している「英語力」（もしくは「英語コミュニケーション能力」）というものは、実際に目に見えるものではない。そうかと言って、存在しないかと言うとそういうわけではない。なぜなら、その能力がない限り英語を話したり、聞いて理解したりすることは不可能だからである。目には見えないけど存在するはずの能力を測定するためには、何らかの方法で児童生徒からそれを引き出して測定できるもの（サンプル）に変えないといけない。それが「テスト」の役目である。一方「評価」は、児童生徒からテストなどを通して引き出したサンプルに価値判断を加えることである。つまり、「テスト」は「評価」の一部である。

児童生徒の能力のサンプルを収集する場合、まず2種類の測定方法（直接テストと間接テスト）に大別できる。そして、収集したサンプルが直接的に評価できるもの（直接テスト）か、引き出したサンプルから推測して間接的に評価できるもの（間接テ

スト）か、どちらかを選択することとなる。どのテストを選択するにしても、適切な評価のために考慮すべき4つの条件（妥当性・信頼性・実用性・波及効果）がある。

##### 4.1. 妥当性

最初に備えるべき条件は、妥当性である。妥当性とは、「テストが測定すべきものを測定しているか」（望月，2015年，p.26）どうかである。それ故に、どのような能力を測るのかをテスト作成時に明確にしておかなければならない。この測定すべき能力が何なのかははっきりしていない限り、妥当な評価はできない。「～できる」というCAN-DOリスト形式の目標の達成を適切に評価するには「～できるかどうか」が評価規準となる。例えば、中学校学習指導要領「外国語」の「話すこと」（やり取り）の目標イでは、「日常的な話題について、事実や自分の考え、気持ちなどを整理し、簡単な語句や文を用いて伝えたり、相手からの質問に答えたりすることができるようにする。」とあるが、この目標を達成できたかどうか評価するためには、実際に上記の課題が「できるかどうか」を測定する必要がある。

では、上述した直接テストと間接テストのどちらの妥当性が高いかという点、直接テストということになる。例えば、英単語の発音の技能を測るために、児童生徒に単語を実際に発音させて評価するのと、紙に記載した単語の正しい発音記号を4つの選択肢の中から児童生徒に選ばせるのでは、どちらが妥当性が高いだろうか。当然ながら、実際に発音させる方が妥当性が高い。なぜなら、実際発音できることが、まさにそのテストで測りたい能力（内容妥当性）であり、それを直接観察して評価できるからである。しかし後者においては、児童生徒は正答が分からなくても正しい選択肢を偶然に選ぶ可能性を否定できない。また、仮に正しい選択肢を選んだとしても、だからと言って正しく発音できるとは限らないからである。こういうわけで、測るべき能力を



明確にし、その能力を引き出す課題を与えて、そのサンプルを直接評価できるプロセスが「妥当的な」評価方法である。

測りたい能力を実際に測定するテストは直接テストが妥当であると上述したが、その内の1つが「パフォーマンステスト」である。パフォーマンステストとは、「話したり、書いたりといったパフォーマンスをさせることによって測りたい能力を直接評価する」（斉田，2019年，p.184）テストのことである。昨今の英語教育では、パフォーマンステストの重要性が高まっている。文部科学省（2017年）は「外国語ワーキンググループにおける審議のとりまとめ」で「学びの過程を通じて、筆記テストのみならず、面接、エッセイ、スピーチ等のパフォーマンス評価、活動の観察等の多様な評価方法から、その場面における児童生徒の学習状況を的確に評価できる方法を選択して評価することが重要である」と述べている（pp.10-11）。児童生徒の学習状況をより的確に適切に評価するためにはパフォーマンステストを適宜活用すべきである。

パフォーマンステストは妥当性の観点から言えば、妥当性の高いテストである。なぜなら、テストが測りたい能力を実際に児童生徒がやってみせるテストであるからである。これを内容妥当性と言う。Brown & Abeywickrama（2019年）によれば、「教室内で内容妥当性を確保する最も実行可能な経験則は、直接パフォーマンスを評価すること」である（p.34：著者訳）。

#### 4.2. 信頼性

テストが備えるべき2つ目の条件は信頼性である。信頼性とは、評価の結果導き出された得点が児童生徒の能力を反映するものだとどの程度信頼できるかである。児童生徒の能力が変わっていなければ、テストにおいてその児童生徒の得点はいつでも同じでなければならない。つまり、テストが測ろうとしている能力以外の要素が得点に一切影響を与え

ていないことが信頼性の高いテストである。児童生徒の英語力以外にテストの得点に影響を及ぼす要素はいつか考えられるが、まず1つ目は評価者である。評価者が2人以上存在する場合、評価者間で評価が変わる可能性が介入する。評価者が2人以上の場合、評価トレーニングを行う必要がある。トレーニングを通して、評価者間の誤差を少なくするのである。望月（2015年）によれば、評価者間の相関を算出し、「相関係数が高ければ（0.7以上）評価者間信頼性は高い」（p.26）ということになる。また、評価者が例え1人であっても、その評価者内で評価がずれることも考えられる。例えば、何十枚、何百枚の英文エッセイを採点しているときに少しずつ基準がずれていくことはありうることである。これでは、評価者内信頼性は高いと言えない。

「評価者間信頼性」と「評価者内信頼性」の両方において、その信頼性を高める1つの手段がルーブリックの活用である。ルーブリックとは、評価の観点（評価基準）と判定基準、その判定基準における児童生徒の学習状況（評価規準）などからなる採点基準表である。このルーブリックを活用することによって、評価者ははっきりと明示された基準に従って評価を行うことができ、評価者が複数いても客観性を担保できる。評価者が1人であっても、ルーブリックに従って、評価を行うことでおれることなく一貫した評価を行うことが可能となる。ルーブリックを綿密に作成することは時間がかかるかもしれない。しかし、その手間をかけることで、評価に迷うことが少なくなり、一貫した評価をおこなうことが可能となり、信頼性の高い評価をおこなうことができる。

#### 4.3. 実用性

3つ目の条件は実用性である。どんなに妥当性と信頼性の高いテストであっても、作成、実施、採点をする際に実用的でなければ、実際にテストは実施できない。実施しようとするテストは、教師の活用

できるさまざまな資源の中で行う必要がある。例えばクラス 40 人全員にスピーキングテストを実施する場合、妥当性を追求して 1 人 1 時間のテスト時間をとるとする。そして、その中で幾つかの話題に関して教師と児童生徒間のインタビュー形式とするとどうだろうか。妥当性という観点からすれば、スピーキング能力を評価するために児童生徒から十分なスピーキングの量を確保することは間違っていない。また、話題によってスピーキングの量も内容も左右されることを考えれば、テスト内の話題も 1 つや 2 つより、もっと多い方がより妥当性の高い評価が下させるはずである。信頼性の観点からも、スピーキングテストの際にもう 1 人の教師と 2 人で評価するとしたらどうだろうか。当然、評価者は 1 人よりも複数のほうがより客観性が担保され、共同で得られた評価はより信頼性が高いであろう。しかし、このようなテストがどんなに妥当性と信頼性が高くても、実用的でなければ実際に実施することは難しい。教師は 1 つのテストに確保できる時間や人材、予算などを十分に勘案しなくてはならない (Brown & Abeywickrama, 2019 年)。実用性の観点からすれば、実用性の高いテストとしてマークシート式テストなどが挙げられる。マークシート式は採点がぶれないという観点から信頼性も高い。しかし、選択肢の中から 1 つ解答を選ぶことで、テストが測りたい児童生徒の能力を適切に測っているかどうかは、既に上述したように吟味が必要である。このように、妥当性と信頼性、そして実用性の間にはトレードオフの関係がある。

#### 4.4. 波及効果

波及効果は、別の言葉で言い換えれば、テストの持つ潜在的な影響力ということができる。テストの影響力の大きさを否定する教師はほとんどいないだろう。児童生徒を勉強に真剣に取り組ませようとするれば、テストを利用する教師はきっと多いはずである。「ここはテストに出やすい」と教師が一言言え

ば、それまで授業に集中していなかった児童生徒でも瞬時に目の色を変えて真剣に取り組み始めることはよくある風景である。つまり、これは一部の児童生徒に当てはまることではなく、一般的に多くの児童生徒に当てはまる現象だからである。金子 (2018 年) は、東京都足立区の中学生約 2,500 人に英語学習に関する意識調査を実施した。金子は、中学生を英語に関する「得意度」と「好意度」に関する質問の回答から、「得意好き群」(n=931 人)、「得意嫌い群」(n=106 人)、「苦手嫌い群」(n=842 人)、そして「苦手好き群」(n=453 人) の 4 つにグループ化した。そして、英語学習の動機づけの要因を特定するための 23 個の動機を提示して、当てはまるかどうかの質問をした。その結果、全生徒の 88% は「英語のテストでいい点数をとりたい」という動機に対して、「よく当てはまる」か「少し当てはまる」と回答している。さらに金子は、上記の 4 グループと 23 個の動機のクロス分析を行ったところ、「英語のテストでいい点数をとりたい」という動機に対して、全てのグループにおいても「よく当てはまる」と「少し当てはまる」が最も多い回答だとわかった。つまり、英語が好きでも嫌いでも、得意であっても不得意であっても、中学生は 10 人中ほぼ 9 人の割合で「英語のテストでいい点数をとりたい」のである。

このようにテストは大きな影響力を持っていることが見てくるわけだが、この影響力は「正」にも「負」にもなりうるのである。Messick (1996 年) は、テストの持つ影響力を英語学習の「促進 (promotion)」と「阻害 (inhibition)」に結び付けて、「正の波及効果 (positive washback)」と「負の波及効果 (negative washback)」に言及している。児童生徒がテストのために学習に取り組み、しっかりと準備した結果、英語力の伸長につながった場合、これは「正の波及効果」である。つまり、テストの存在が児童生徒の学習を「促進」して、本来の目的である「英語力」を向上させることにつながっ

たからである。一方、テストの存在が、児童生徒の学習を「阻害」することもある。例えば、テストがいつも文法や単語などの明示的知識のみしか測らないのなら、児童生徒はおそらくテスト対策としてスピーキングやライティング、リスニングなどの練習はおこなわないであろう。その結果、児童生徒の英語の明示的知識は豊富になるが、実際のコミュニケーションに必要なスピーキング力やリスニング力、ライティング力などは十分に育たないと推測できる。つまり、上述したようなテストで高い得点を望む結果、児童生徒は明示的知識のインプットに時間を費やし、日本の英語教育が本来育てようとしている「英語力」の伸長につながらないのである。さらには、英語の文法知識などを重視したテストが英語の文法中心の授業展開を招くことになり、「英語嫌い」を生み出すことにつながるのである。上述の金子の調査では、「英語嫌い」を生み出す最大の要因は「文法」である。「英語嫌い群」の73%が「文法が難しい」と回答している。もし授業が「文法中心」で展開されていけば、「英語嫌い」が誘発され、多くの児童生徒の英語学習が「阻害」されるのである。これが、「負」の波及効果である。こういうわけで、テストは「正の波及効果」にも「負の波及効果」にもなりうることを常に意識して、正の波及効果を生み出すテストの作成を考えなければならない。

## 5. 理論と実践の融合

### 5.1. 目的

上述した理論的背景を基に、指導と評価の一体化の実践を検定教科書を使って例示する。指標形式の目標を設定し、目標達成を評価するための評価方法を具体的に提示することで、学校現場で指導にあたる教員の一助としたい。本稿で提示する例は、実際の学校での実践を想定して作成されているものである。そのため、これを実際に実践してもらうことを願う。もちろんこのまま実践してもらってもよい

し、実際のクラスに合わせて修正して使ってもらってもよい。

### 5.2. 指標形式の目標設定からの逆向きデザイン

本稿では、中学校2年の教科書「Columbus 21 年」(光村図書)のUnit 3を取り上げる。この単位では、新出文法項目として、“be going to do”や“will”が導入される。これらの表現を使って、夏休みの予定を登場人物たちがたずねあう内容になっている。そこで、「未来の予定をたずねあうことができるようにする」という指標形式の目標をこの単位の目標の一つとして設定する。

次に、上記の目標の達成度を評価する評価方法を考える。評価方法というと一連の指導後におこなうもので最後に実施することになることが多いが、その評価方法を単位の最初の計画の段階で決めておくことが肝要である。なぜならば、指導の結果として、児童生徒にどのような学習状況が見られるかを教師がはっきり思い描いて指導に臨むことができるようにするためである。このように指導計画を考える際に、児童生徒に達成させたい目標を設定してその目標の達成度を測る評価方法も最初に決定しておき、最後の指導結果から逆算して指導計画を考えいくプロセスは「逆向きデザイン」(Backward Design)と呼ばれている。逆向きデザインを活用して、上記の単目標を適切に評価する方法を考える。評価方法は目標によって変わる。既に上述したように、適切な評価方法を選択することが大切であるが、その際には上述の4つの条件(妥当性・信頼性・実用性・波及効果)を十分に考量する必要がある。

### 5.3. 適切な評価方法の選択

ここからは、上記の4つの条件を満たす適切な評価方法を実際に考えていく。まず妥当性の観点から見えていくと、上述したように「測りたい能力を実際に測っているか」が問われる。ここでの目標は「未



来の予定をたずねあうことができるようにする」ことなので、実際に生徒が「未来の予定をたずねあうことができるかどうか」を判断する必要がある。そのために最も妥当な評価方法は、生徒が実際に「できるかどうか」を引き出して、その結果を評価することである。つまり、パフォーマンステストが最も妥当な評価方法だと言える。

パフォーマンステストの場合、測りたい能力をどのように引き出すかが大切なポイントである。そのために、測りたい能力を引き出すための適切な「課題」の設定が必要不可欠である。もし適切な課題が設定されなければ、測りたい能力を引き出せずに妥当な評価が下せない。パフォーマンステストの課題を考える場合、できるだけ具体的な場面設定が望ましい。ここでは「未来の予定」をより具体的に「週末の予定」をたずねあうという課題とし、生徒Aと生徒Bの2人の「やり取り」を設定する。ここで引き出したい能力は「たずねあうこと」なので、当然生徒は誰かと会話をしなくてはならない。しかも、生徒はたずねるだけでなく、自らも予定を伝えることができないといけない。教師と生徒のやり取りも考えられるかもしれないが、ここでは生徒同士にやり取りをしてもらうこととする。なぜなら、実用性の観点も考慮すべきだからである。生徒2人のやり取りなら、単純に考えても、教師と生徒の会話の半分の時間で実施できることとなる。学校現場の限られた時間を考えれば、「生徒2人によるやり取り」とした方がより実用的である。

次に、評価者をどうするかである。この場合、教師が自ら評価者となることがもっとも実用的である。もしこのパフォーマンステストを複数のクラスや学年全体で実施できるならば、同僚の英語教員と協力しておこなうこともできる。いずれにしても、評価をする当事者は1人と考えるのが实际的である。2人で評価をおこなうことは、学校現場にALTがいたり、相当教員の数に余裕があったりしない限り難しい。ただし、パフォーマンステストの場には

評価者は1人かもしれないが、ビデオ撮影をしておくことで、別な英語教員に評価をしてもらうことも可能である。そうであれば、同時に評価をする必要はなく、実現可能性は高くなり、信頼性もより高くなる。

信頼性の観点で大切なことは、評価基準がぶれないことである。評価者が複数いる場合は尚更であり、1人であっても基準がぶれることは考えられる。この評価基準のぶれを防ぐために、上述したようにルブリックの作成を考えるべきである。ルブリックの作成においていくつか考慮する点がある。まず、評価の基準の明確化である。評価の基準とは、言い換えれば、「測りたい能力」である。ここで測りたい能力は、「週末の予定をたずねあう」能力である。アウトプットのクオリティを評価する際、用いられる基準が3つあり、「正確さ」(accuracy)と「流暢さ」(fluency)と「複雑さ」(complexity)である(Skehan & Foster, 1996; Yuan & Ellis, 2003)。「複雑さ」とは、文法の高度さや文構造の複雑さのことである。本稿で取り上げているパフォーマンステストにおいて、複雑な文構造や高度な文法は要求されているとは言えない。そのため、「複雑さ」の観点はここでの評価基準に含めないこととし、「正確さ」と「流暢さ」を主な基準とする。

#### 5.4. パフォーマンス・モデルの作成

正確さと流暢さも幾つかの下位能力がある。例えば、正確さには、文法や語彙使用の正確さや語用的な正確さ、発音やアクセントなど音声面での正確さなどに分類できる。「流暢さ」は内容を中心とした発話量や発話のスピードなどに分類できる。このように評価できる観点はたくさんある。しかし、全ての観点を毎回評価することは時間も労力もかかり、妥当とも言えない。むしろ、パフォーマンステストごとに観点を焦点化して評価すべきである。そのためには、パフォーマンスモデルを作成することを勧めたい。このモデルを作成することで、実際の生徒



に育成したい能力が明確化してくる。以下に、そのモデルを提示する。

A: Hi, B. How are you?  
 B: Hi, A. I'm good. Thank you. How are you?  
 A: I'm fine, too. Thank you.  
 Do you have any plans for the weekend?  
 B: Yes, I'm going to see a movie with my friend.  
 How about you? Do you have any plans for the weekend?  
 A: I'm going to practice soccer. Nice talking with you. See you.  
 B: Nice talking with you, too. See you.

当然のこととして、あいさつをすることを身につけさせたい。日常生活において、いきなり質問から会話を始めることは「まれ」な状況だ。実際のコミュニケーションを意識した真正性（authenticity）の観点からも、必ず一連のあいさつで会話を始め、閉じるように指導をしたい。そのために、上記のモデルに示している。次に、「未来の予定についてたずねあう」能力を評価したいので、生徒 A は “Do you have any plans for the weekend?” などとたずねることが要求される。生徒 B は “be going to” を使って自身の週末の予定を伝えることが求められる。生徒 A と生徒 B は役割を交代して同様のやり取りが求められる。

学校現場では、生徒の実情に応じて「学習到達目標」を独自に設定することが求められている。よって、学校ごとに指導の目標が変わり、生徒に求めるレベルも変わってくる。上記に示したモデルは「課題」で求めている基本レベルと言える。学校によっては、より発展的なレベルを生徒に求めていけるはずだ。そこでより発展的なパフォーマンスモデルを以下に例示する。

A: Hi, B. How are you?  
 B: Hi, A. I'm good. Thank you. How are you?  
 A: I'm fine, too. Thank you.  
 Do you have any plans for the weekend?  
 B: Yes, I'm going to see a movie with my friend. I'm so excited.  
 A: Wow, sounds great. Which movie are you going to see?  
 B: We are going to see “Jurassic World.”  
 How about you? Do you have any plans for the weekend?  
 A: I'm going to practice soccer. We have a big tournament next weekend.  
 B: I see. Are you excited about the tournament?  
 A: Yes, I am. Nice talking with you. See you.  
 B: Nice talking with you, too. See you.

実際の日常会話で “be going to” を使って自身の予定を伝えて、それで終わってしまうことはまずないはずである。むしろその後に、上記の下線部のように自分の気持ちやプラスアルファの情報など「プラス 1」（plus 1）を付け加えることの方が自然であり望ましい。だから、生徒にはより真正性の高いパフォーマンスを求めたい。さらに、それを聞いた相手も回答を聞いた後に何も言わないのは不自然であり、日常生活のコミュニケーションでは大きな問題となる。回答をもらった後、それに対してつなぎ言葉を用いたりして「フィードバック」（feedback）してほしいものである。そして、会話を発展させるための「追加質問」（follow-up question）も求めたい。よって、「未来の予定をたずねあうことができるようにする」という目標から、「聞き手に配慮しながら、未来の予定をたずねあい、対話を続けることができるようにする」という目標にレベルアップさせる。それに対応して、パフォーマンステストの課題を「週末の予定をたずねあう」から、「聞き手

に配慮しながら、週末の予定をたずねあう。そして、つなぎ言葉を用いて、対話を続ける」という課題に変更する。「聞き手に配慮しながら」は今回の学習指導要領改訂で初めて示された用語である。小学校外国語活動での「相手への配慮」や小学校外国語での「他者への配慮」と接続するキーワードとなっている。中学校では、常に「聞き手（読み手、話し手、書き手）」に配慮しながらコミュニケーションを図ることは大切である。このように教師自らがパフォーマンスモデルを作成・実演してみると、生徒に何を求めるかが明確になる。生徒へ求めるレベルがはっきりしたことで、この単元の目標もそれに対応させる形に変更することができる。

### 5.5. ルーブリックの作成

上記のパフォーマンスモデルの作成を経て、生徒に求める能力が具体化できた。ここで、ルーブリックを作成するための評価の観点を整理する。まず、「週末の予定をたずねあう」能力を測ることから、“be going to”を使って未来の予定を表す発言は必ず評価したい項目である。また、「たずねあう」ことを要求しているので、“be going to”を肯定文でも、疑問文でも正確に使ってほしい。そこで、ターゲット文法である“be going to”を含む正しい文の使い方が焦点となる。この観点を「語法」と言うことができる。そして、この観点は新学習指導要領における評価の3つの観点のうちの1つである「知識及び技能」に当てはまるものである。

次に「流暢さ」の観点であるが、上記のパフォーマンスモデルで示したように、流暢に会話をつなげるためにも相手の発言に対して反応してほしいし、会話を発展させる追加質問もしてほしいので評価に取り入れたい。そして、追加質問に対しても、適切に返答してほしい。これら一連のやり取りを「内容」と言うことができる。そして、この観点は、上記に上げた新学習指導要領における評価の3観点の「思考力・判断力・表現力」に当てはまるものであ

る。文部科学省（2017年）は、「自分の…（中略）…考えをただ一方的に伝えるのではなく、相手が話したり…（中略）…した内容にも十分に注意を傾けながらやり取りをし、お互いの理解を深められるようにしていくことが重要」（pp. 53-54）としており、まさにこの能力をこのテストで評価している。

最後に、スピーキングテストであるので声の大きさや明瞭さも重要である。相手に聞こえなければ、いくら正確な英語を話しても伝わらないし、発音が不明瞭でも、やはり伝わらない。しかし、声の大きさと発音を別々に評価するのは評価の項目が多くなり、評価者にとって負担となる。この2つは一体化して評価したい。そこでこれら2つを「声の明瞭さ」とする。また、アイコンタクトもコミュニケーションにおける大切な要素である。よって、「声の明瞭さ」と「アイコンタクト」は、コミュニケーションを成立させるための「主体性」としてまとめて評価したい。そして、この観点も新学習指導要領における評価の3観点の最後の観点である「学びに向かう力・人間性等」に当てはまる。なぜなら、上述の文部科学省は、「主体的に外国語を用いてコミュニケーションを図ろうとする態度」（p. 16）をこの観点で評価することを例示している。

上述の3つの評価の観点を一番左の縦列に表記し、横列にその観点ごとの生徒のパフォーマンスの評価規準を配点ごとに配列して、ルーブリック（表1）を以下に作成した。

この単位では、“be going to”を使って未来の予定を伝え合う活動だけでなく、会話の始め方や終わり方、自分の意見にプラス1の情報提供、相手の発言に対してのフィードバックや追加質問など、会話の自然な発展のさせ方を指導する必要がある。パフォーマンステストを実施する前に、これらのことを扱った言語活動に授業の中で取り組ませる必要がある。当然ながら、指導していないことを生徒に求めることはできないからである。上記のルーブリックが完成すれば、この単元の終わりに生徒が英語を

表1 ルーブリック

	A (5点)	B (3点)	C (1点)
語法 (知識・技能)	be going to を正確に使って相手の予定をたずね、自分の予定を伝えている。	左記のどちらか一方が正確に言えている。	左記の両方とも正確に言えない。
内容 (思考力・判断力・表現力)	プラス1やフィードバック、追加質問を全て満足にできている。	プラス1やフィードバック、追加質問の3つの内2つは満足にできている。	プラス1やフィードバック、追加質問の3つの内1つか全く満足にできていない。
主体性 (学びに向かう力・人間性等)	アイコンタクトと声の明瞭さは十分である。	アイコンタクトと声の明瞭さの内、1つは十分である。	アイコンタクトと声の明瞭さの両方とも不十分である。

使って何ができるようになっていくかを具体的にイメージすることができる。このイメージを持つことで、教師は単元を通して指導すべきことがはっきりと見えてくる。教科書だけ扱っては見えてこない内容である。指導の目標とその評価法を明確にしてこそ、指導内容が鮮明になるのである。

## 6. 結 論

本稿では、昨今の英語教育の課題の1つとなっている「指導と評価の一体化」に関して、理論的背景を基に、中学校の実際の教室・教科書で実践できる1つのアプローチを具体的に提示した。特にルーブリックの作成においては、パフォーマンスモデルを作成することで評価すべき観点を焦点化させ、具体化させる手法を提示した。ルーブリックの作成に関しては、決まったルールは存在しないため、教師の裁量に任されている。本稿では、評価法の理論を基にルーブリック作成の1つの方法を提示した。これによって、児童生徒のどんな英語力を評価すればよいかを具体的に考察した。そして、評価から逆算することで授業で指導すべき内容を明確化して、それを基に指導を行う指導と評価の一体化のアプローチを例示した。新学習指導要領が公示され、全面実施まであとわずかである。英語教育は今後も変化を求められていく。指導と評価の一体化は今後ますます重要になっていく。本稿で示したアプローチが学校

現場の教員の一助となることを願っている。

### 〈引用文献〉

- Brown, H.D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices*. (3rd ed.). Hoboken, NJ: Pearson Education.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.
- Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18: 299-324.
- Yuan, F. & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24a: 1-27.
- 金子 義隆. (2018). 「英語教育を育むための教育的示唆」『明海大学教職課程センター研究紀要』創刊号, 2-14.
- 斉田 智里. (2018). 「学習評価のあり方・進め方：目標に基づく指導と評価の一体化の実現のために」『「学ぶ・教える・考える」ための実践的英語科教育法』東京：大修館書店.
- 高梨 庸雄・卯城 祐司 (編). (1981). 『英語リーディング事典』東京：研究社出版.
- 望月 昭彦. (2015). 「ライティングの評価」『英語4技能評価の理論と実践——CAN-DO・観点別評価から技能統合的活動の評価まで』東京：大修館書店.
- 文部科学省. (2011a). 『国際共通語としての英語力向上のための5つの提言と具体的施策～英語を学ぶ意欲と使う機会の充実を通じた確かなコミュニケーション能力の育成に向けて～』  
[http://www.mext.go.jp/b\\_menu/shingi/chousa/shotou/102/houkoku/1352460.html](http://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/1352460.html)



- 文部科学省国立教育政策研究所. (2011). 『評価規準の作成, 評価方法等の工夫改善のための参考資料 中学校外国語』 研究社.
- 文部科学省初等中等教育局. (2013). 「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」 [http://www.mext.go.jp/a\\_menu/kokusai/gaikokugo/\\_\\_icsFiles/afieldfile/2013/05/08/1332306\\_4.pdf](http://www.mext.go.jp/a_menu/kokusai/gaikokugo/__icsFiles/afieldfile/2013/05/08/1332306_4.pdf).
- 文部科学省. (2014). 「今後の英語教育の改善・充実方策について 報告～グローバル化に対応した英語教育改革の五つの提言～」 [http://www.mext.go.jp/b\\_menu/shingi/chukyo/chukyo4/015/gijiroku/\\_\\_icsFiles/afieldfile/2014/11/11/1353308\\_12.pdf](http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/015/gijiroku/__icsFiles/afieldfile/2014/11/11/1353308_12.pdf).
- 文部科学省. (2016). 「外国語ワーキンググループにおける審議の取りまとめ」 [http://www.mext.go.jp/component/b\\_menu/shingi/toushin/\\_\\_icsFiles/afieldfile/2016/09/12/1377057\\_1\\_1.pdf](http://www.mext.go.jp/component/b_menu/shingi/toushin/__icsFiles/afieldfile/2016/09/12/1377057_1_1.pdf).
- 文部科学省. (2017). 『中学校学習指導要領解説 外国語編 英語編』 東京：開隆堂
- 文部科学省. (2018). 『高等学校学習指導要領解説 外国語編 英語編』 東京：開隆堂